Working with Light

Working with Light for Computer Graphics

Physics and Optics:

- Light is <u>emitted</u> from a <u>light source</u>
  - e.g. the sun, a light bulb, computer monitor, cell phone, etc.
- That emitted light impacts various objects, and may be <u>reflected</u> or <u>absorbed</u>
  - This reflection/absorption modifies the light
    - e.g. creating color, brightness, dullness/shininess, highlights, etc.
- In addition, light may <u>pass (transmit) through</u> various materials and (in doing so) be bent, scattered, etc.
  - e.g. prism, stained glass windows, water, etc.

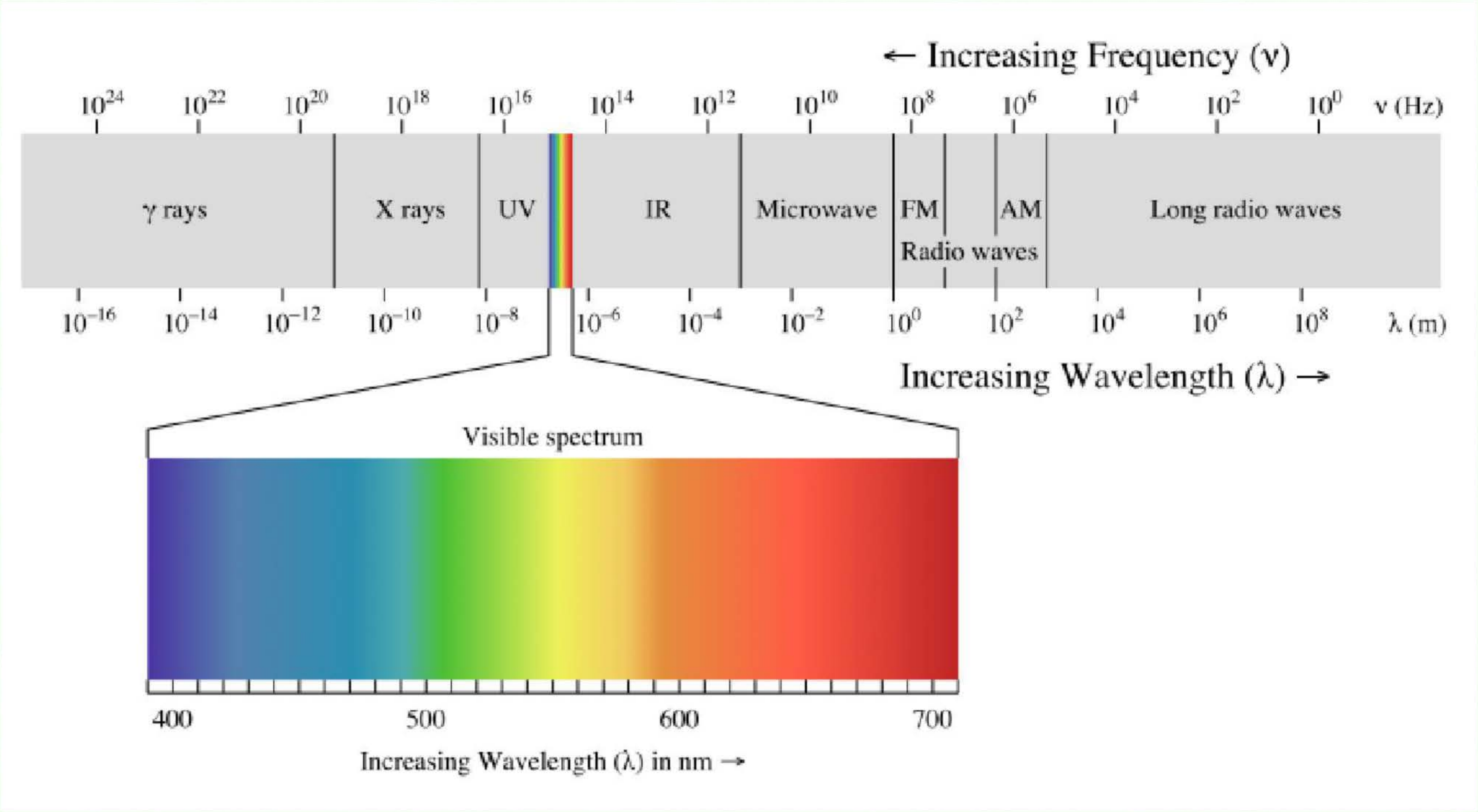
Human Perception:

- Eventually, some light may enter our <u>eyes</u> creating a <u>signal</u>
- Our <u>brain</u> creates an <u>image</u> based on the signals it gets from our eyes

Software/Hardware:

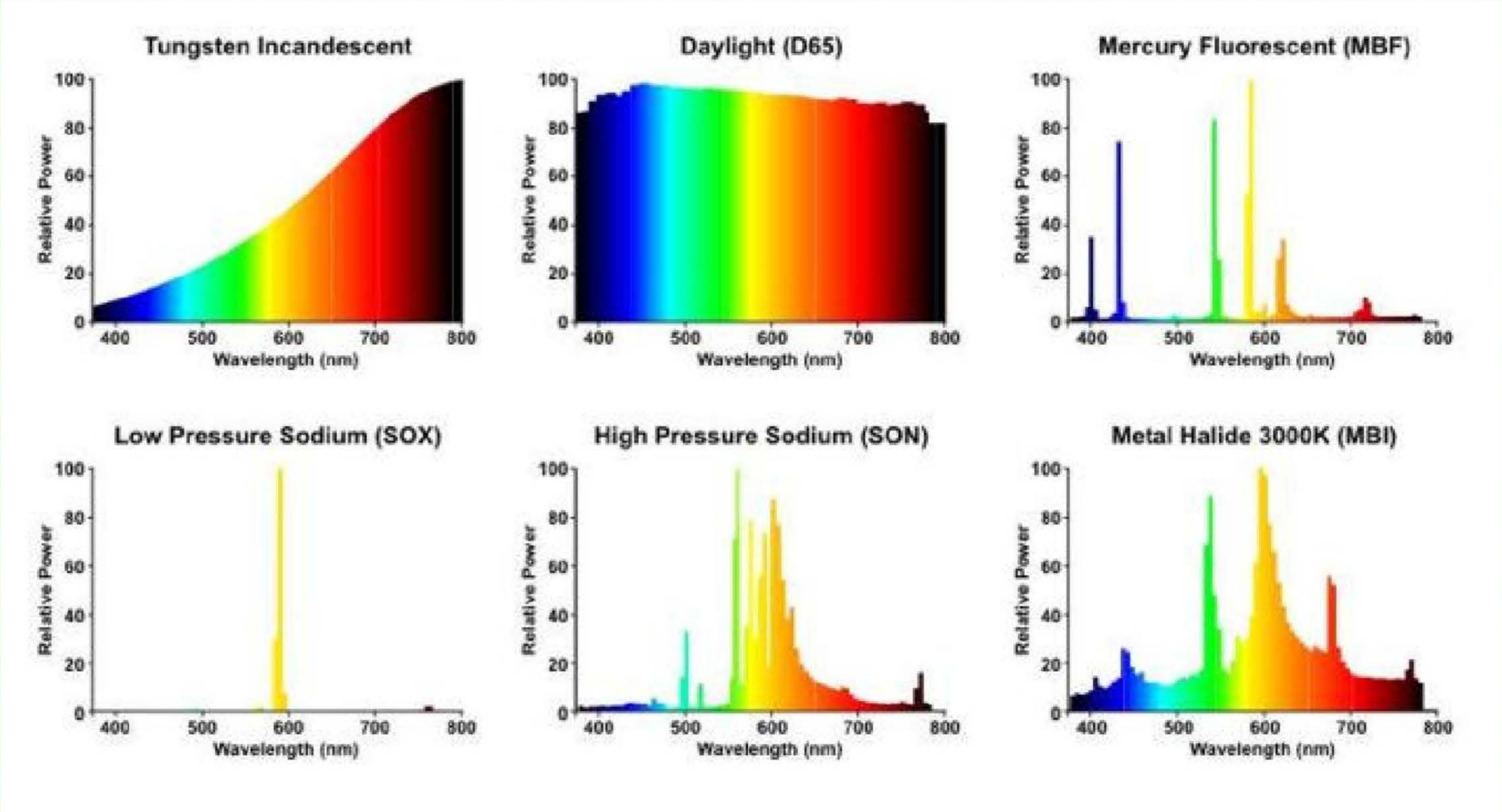
- Understanding <u>the physics of light (i.e. optics)</u> is important
- Understanding human perception allows for MANY <u>optimizations/simplifications</u> in both software/hardware
- The images we create ARE NOT intended to duplicate reality, only to fool humans into believing such

Electromagnetic Spectrum



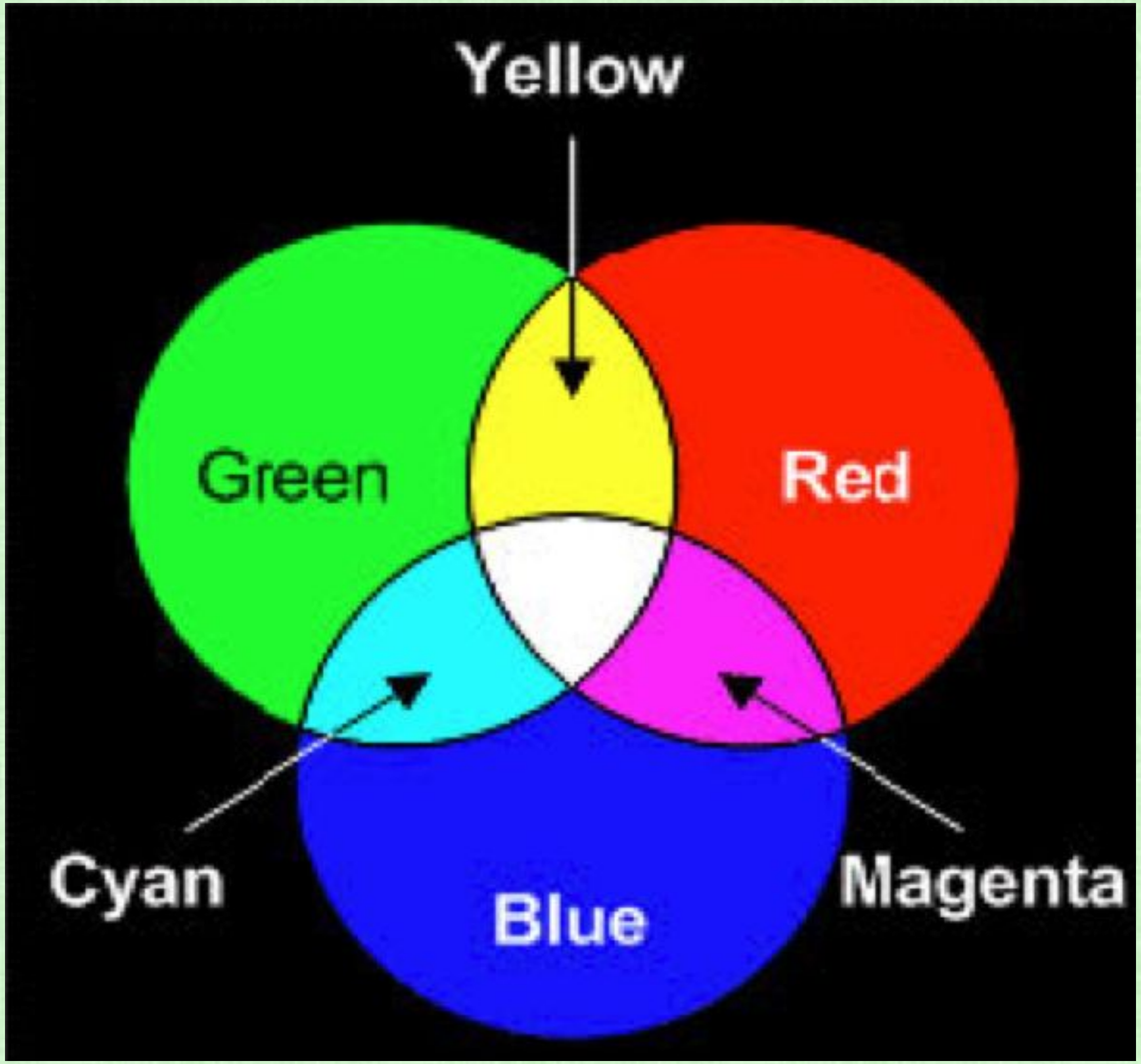
- **The human eye can only see wavelengths between about 400 nm to 700 nm, so we focus on those**

Relative Power Distribution of Lights



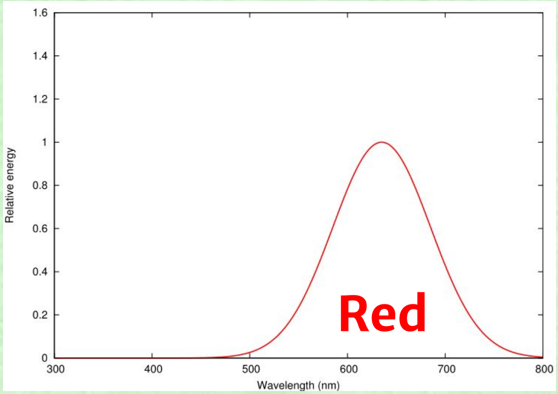
Adding Light Energy

The human eye perceives combinations of light energy as follows:

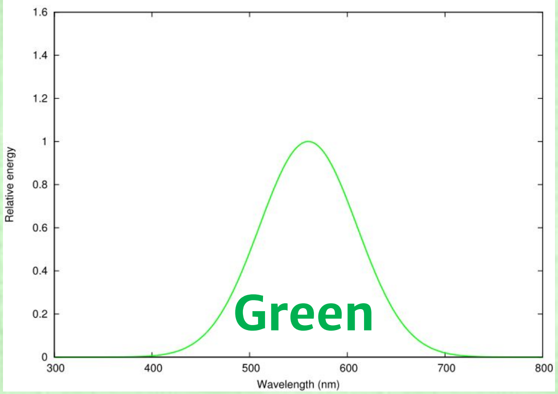


Adding Light Energy

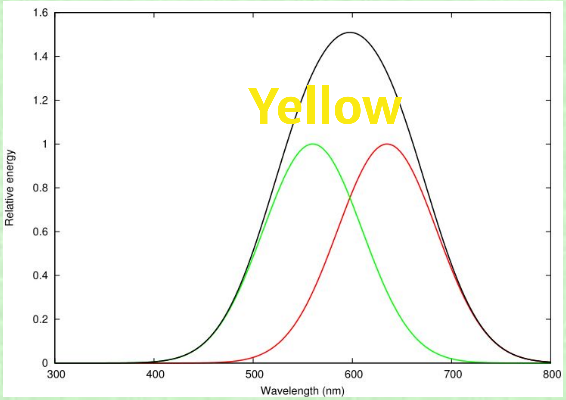
- **Energy adds (per wavelength) according to:** $E(\lambda) = E_1(\lambda) + E_2(\lambda)$
- **This leads to the following relative power distributions:**

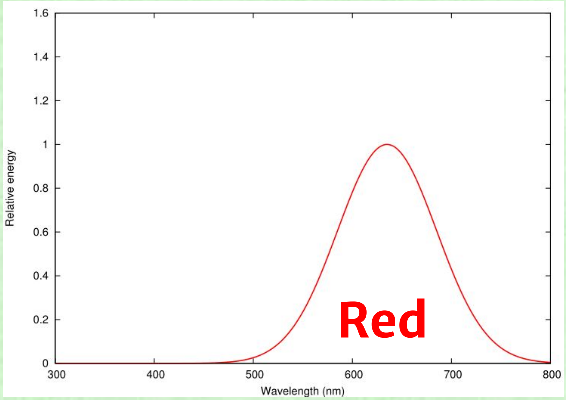


+

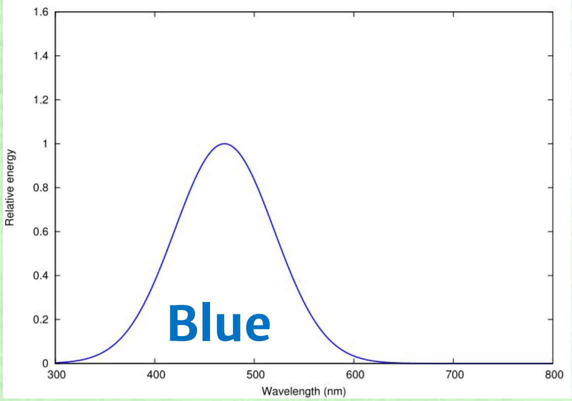


=

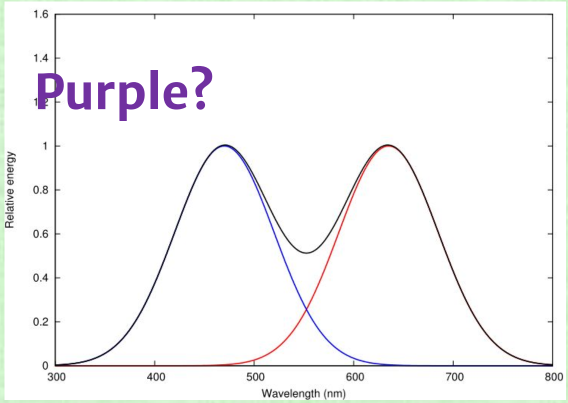




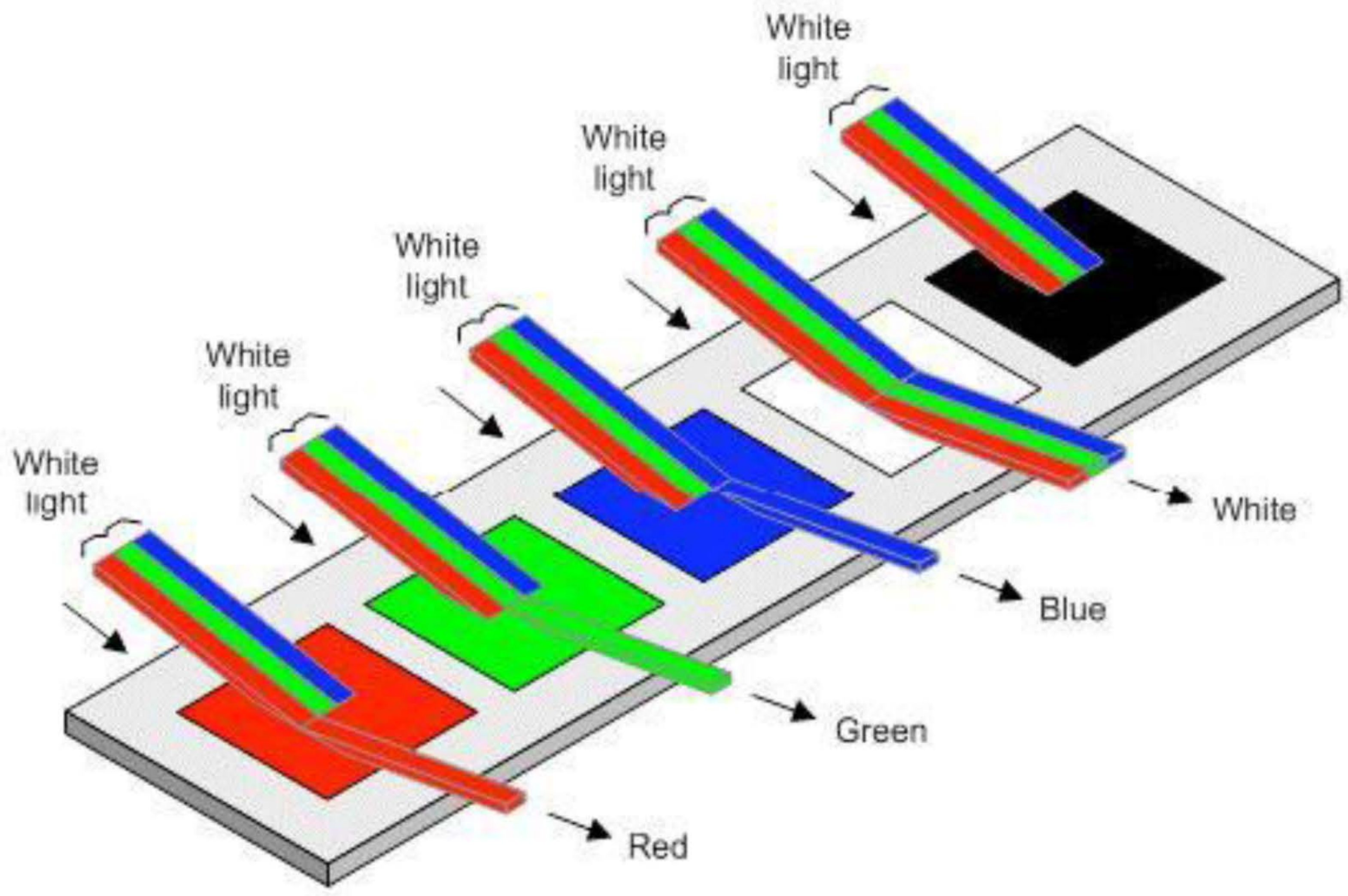
+



=



Absorbing & Reflecting Light Energy



Shining white light on different colored paints

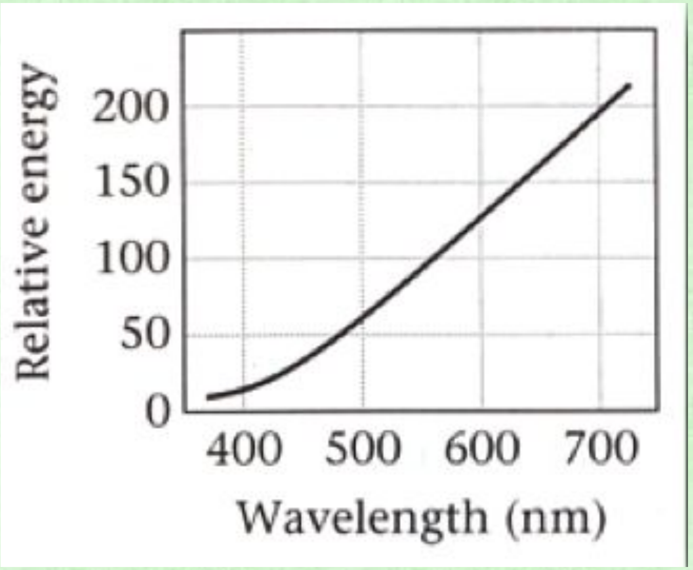
Absorbing & Reflecting Light Energy

All light energy is either reflected or absorbed: $\mathbf{r}(\boldsymbol{\lambda}) + \boldsymbol{a}(\boldsymbol{\lambda}) = \mathbf{1}$

$$\mathbf{0} \leq \mathbf{r}(\boldsymbol{\lambda}), \boldsymbol{a}(\boldsymbol{\lambda}) \leq \mathbf{1}$$

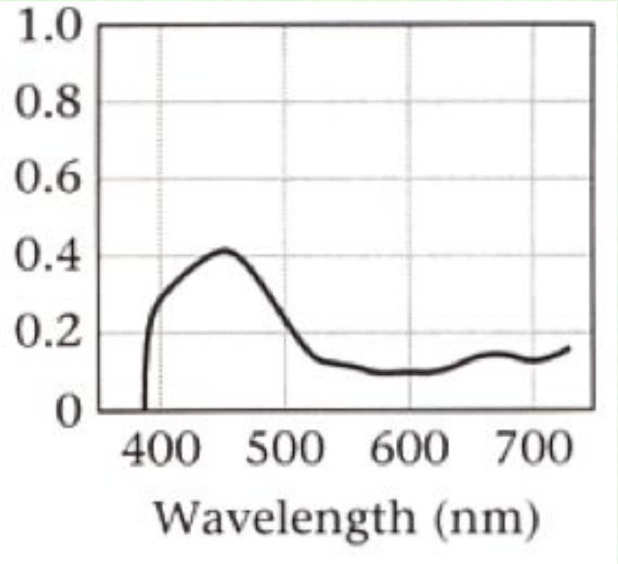
Reflected light energy (per wavelength) is computed via:

$$\boldsymbol{Reflected}(\boldsymbol{\lambda}) = \boldsymbol{E}(\boldsymbol{\lambda})\boldsymbol{r}(\boldsymbol{\lambda}) = \boldsymbol{E}(\boldsymbol{\lambda})\big(\mathbf{1} - \boldsymbol{a}(\boldsymbol{\lambda})\big)$$



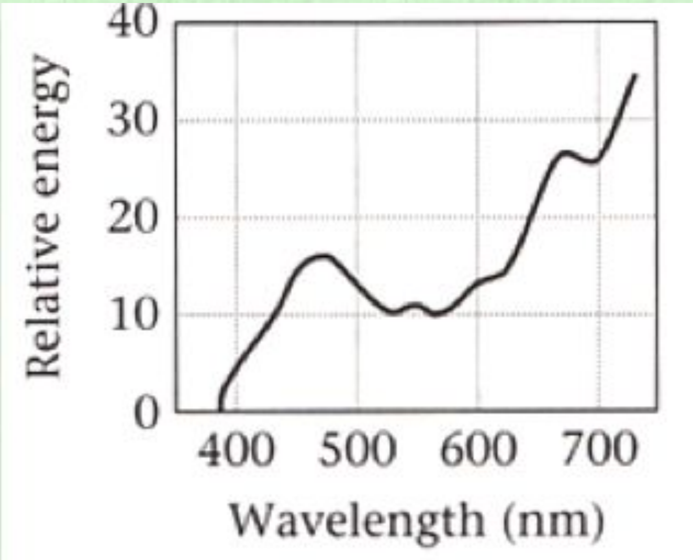
Incoming Energy

×



Material’s Reflectance

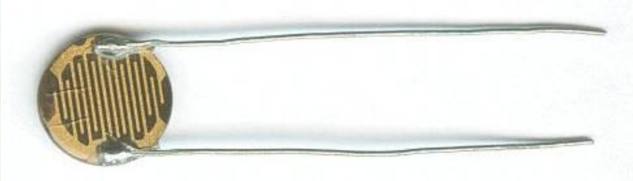
=



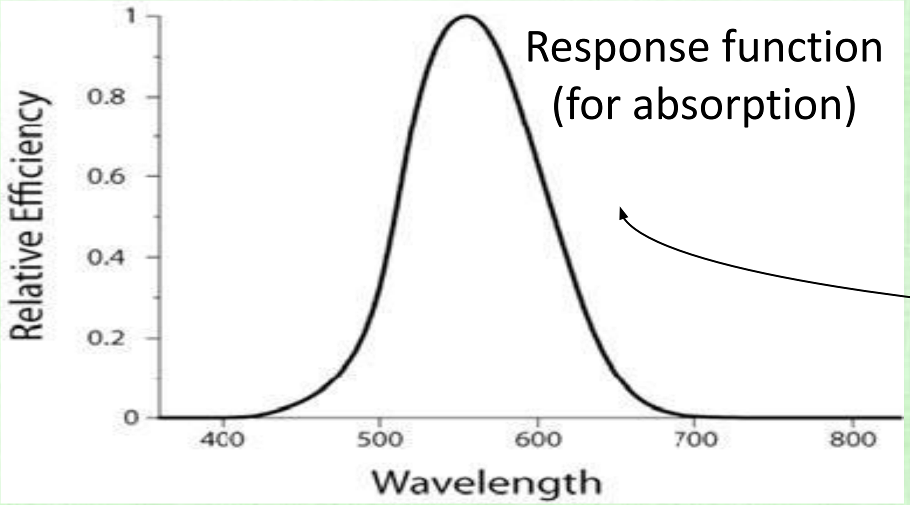
Reflected Energy

Sensor Absorption

- Sensors absorb light (per unit time) and create a signal (per unit time)
- In order to be small (both biologically/mechanically), they are highly specialized



- Specialization leads to sensors only creating one signal (per unit time) for the entire sensor
- They conflate all the various wavelengths of light they absorb into one signal (per unit time)



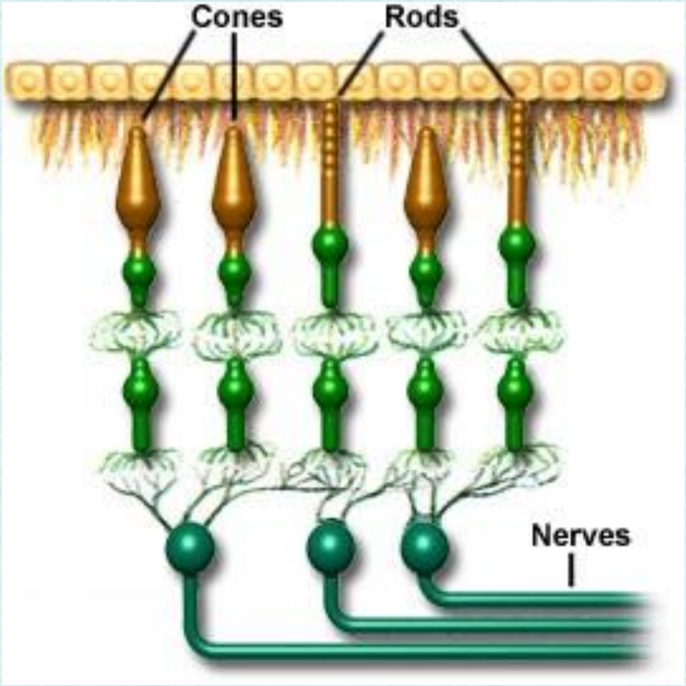
Signal power (energy per second):

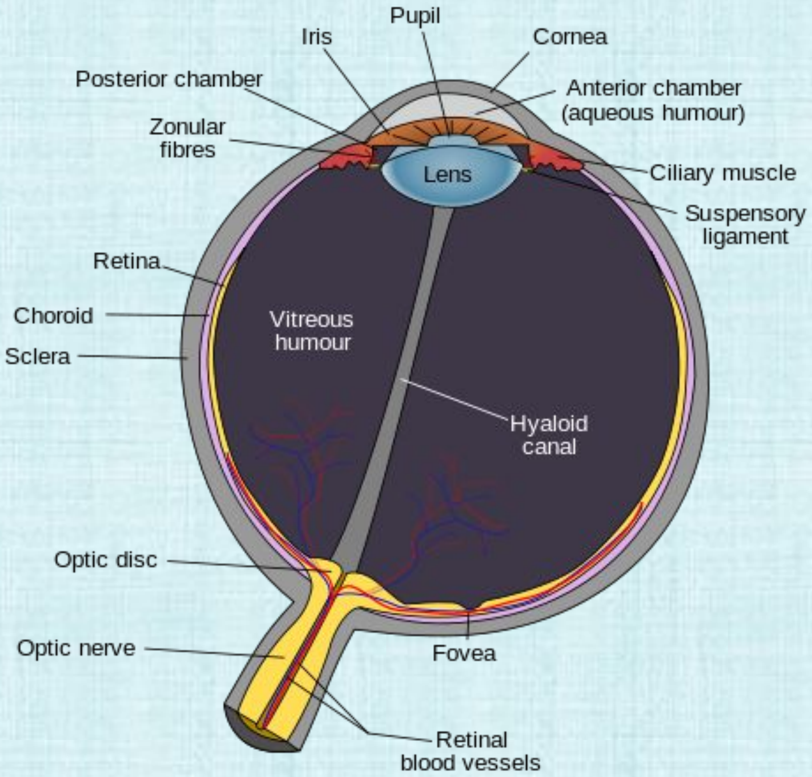
$$A = \int E(\lambda)a(\lambda)d\lambda$$

Not all wavelengths contribute equally to the final signal

Sensors in the Human Eye

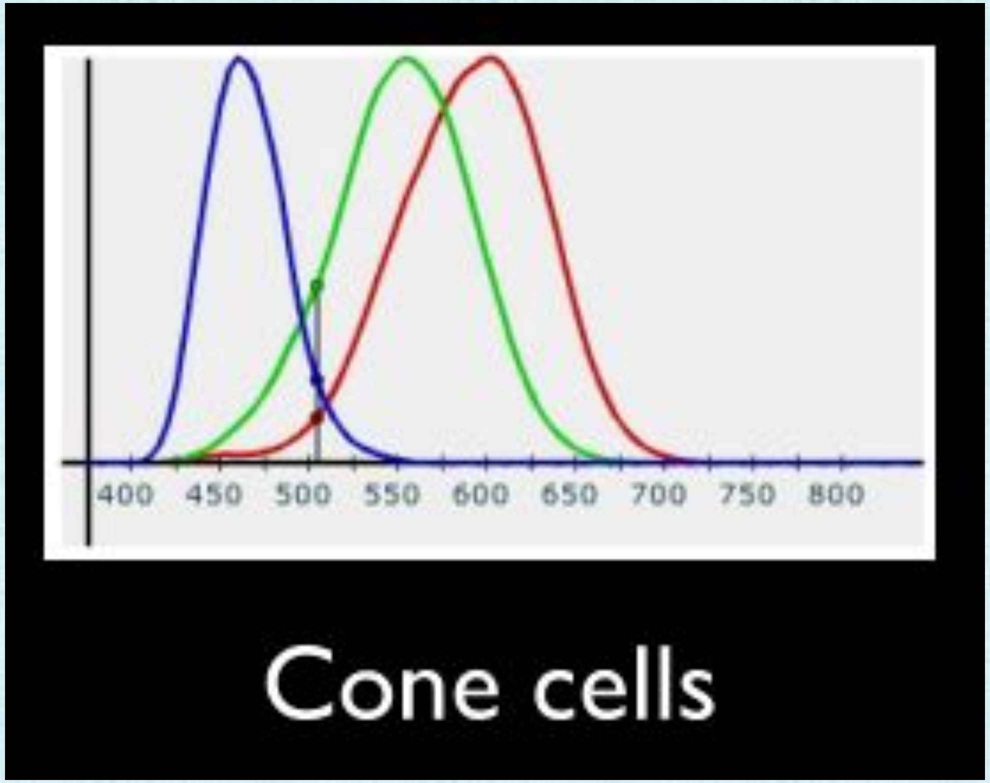
- **The eye has 3 different kinds of cone sensors and 1 rod sensor**
- **Proteins in the cone/rod cells absorb photons changing the cell's membrane potential**
- **At <u>night</u>, cones are under-saturated (no/low/noisy signal), and rods produce most of the understandable signals**
- **During the <u>day</u>, the rod signals are over-saturated (all clamped to max), and we see primarily with our cones**



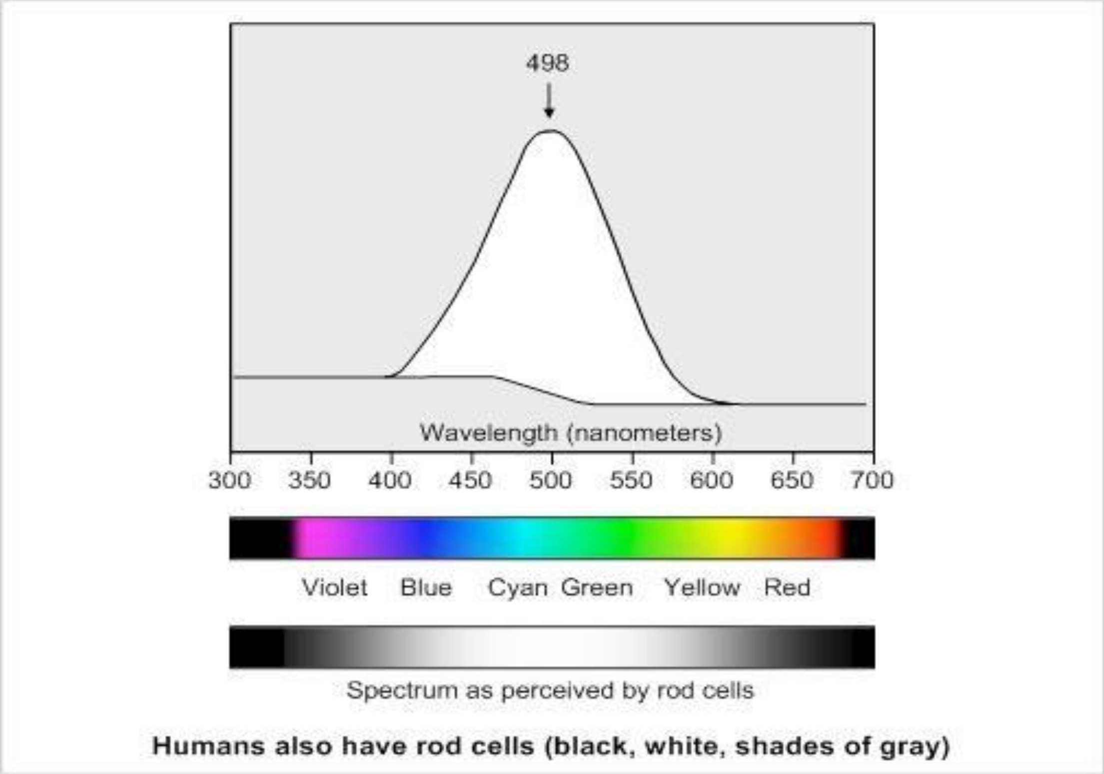


Response Functions for Human Eye Sensors

- **The 3 cone sensors each have response functions focused on different wavelengths, and are referred to as red, green, and blue cones**
- **The single rod sensor is interpreted as a black/white/gray light intensity**



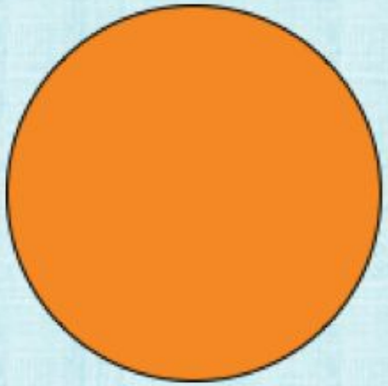
Note the similarity in red/green
(in regard to red/green colorblindness)



At night, the *single* signal from the rod can at best
be understood as a shade of gray

Trichromatic Theory

- Given any human perceived “color”
- Can adjust the brightness of 3 single wavelength lasers (e.g. R = 700 nm, G = 546 nm, B = 435 nm) to fool a human observer into “<u>mistakenly</u>” thinking that the laser combination matches that “color” (i.e. as a single wavelength)
- This is doable because each of the three cones can only send one signal (3 dimensional basis)



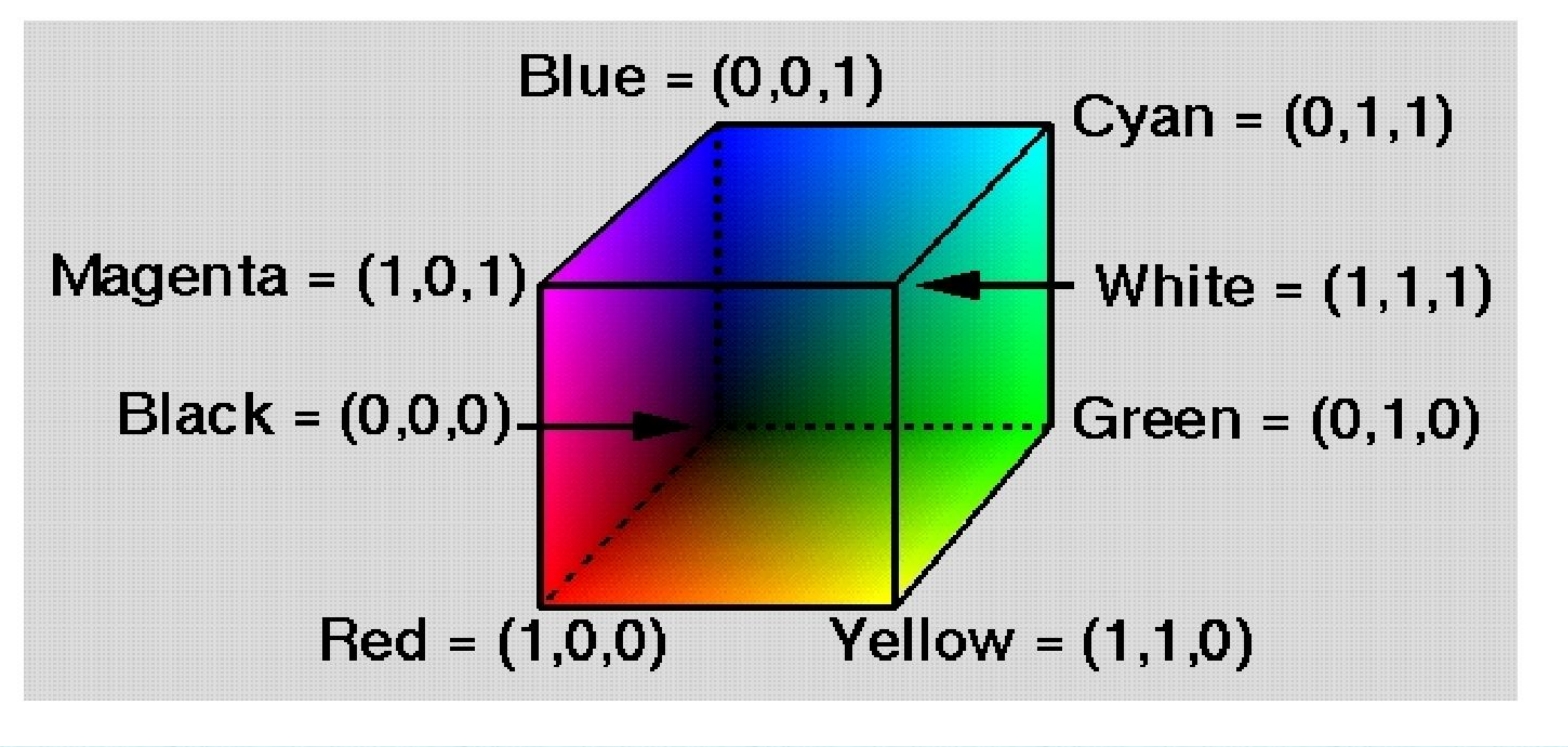
“=
“

C “=“ R “+” G “+” B

- Since the eye only perceives 3 signals (ignoring rods), we only need 3 signals for images, cameras, printers, displays, etc. (human perceived color is a three dimensional space!)
- Image formats store values in the R, G, and B channels

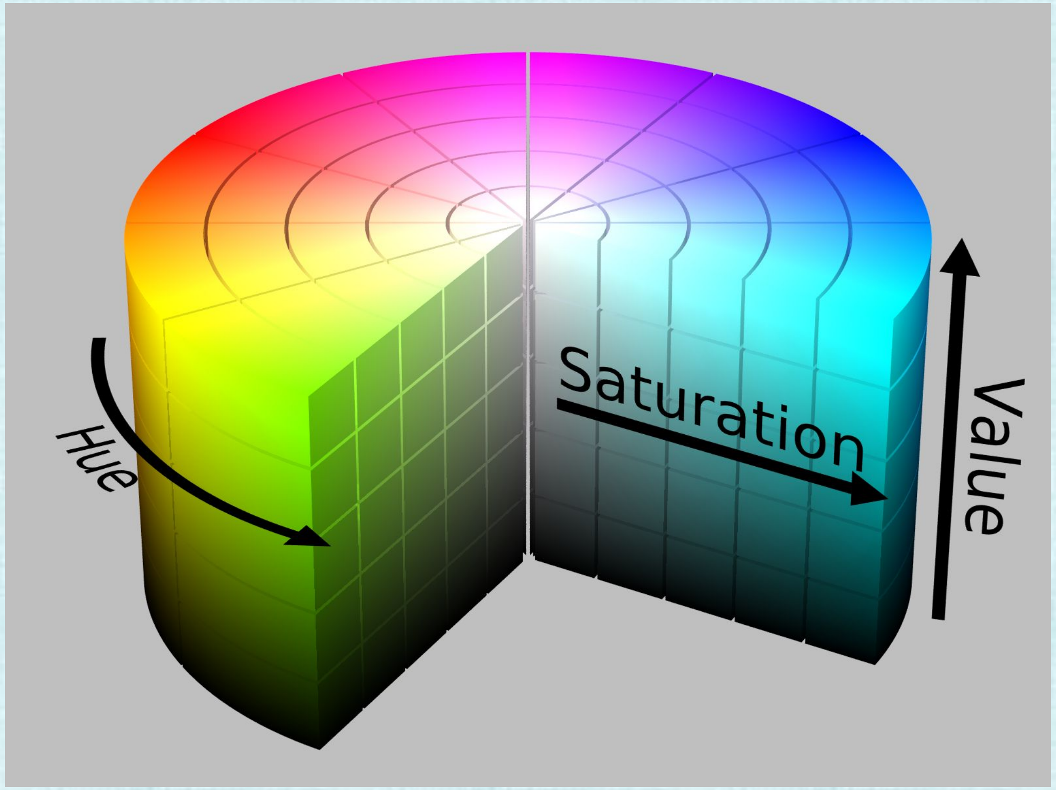
Three-Dimensional Color Space

- **Map each primary color (Red, Green, Blue) to the unit distance along the x, y, z axes**
- **Black at (0,0,0), white at (1,1,1)**
- **The resulting RGB Color Cube represents all possible colors**



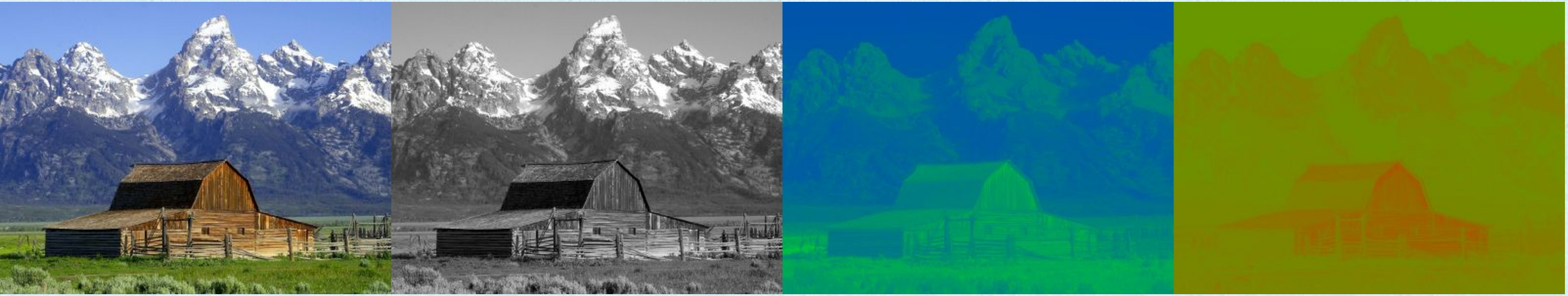
Cylindrical HSV Color Space

- **A better 3D color space for <u>user interfaces</u> is based on Hue, Saturation, and Value (HSV)**
- **<u>Hue:</u> rainbow of colors ("wavelength")**
- **<u>Saturation:</u> intensity for a particular color**
- **<u>Value:</u> lightness or darkness of a particular color**



Luminance and Chrominance (YUV)

- **Another 3D color space represents an RGB color via 1 luminance (Y) and 2 chrominance (UV) channels**
- **Black and White televisions used Y only, which perceptually holds the most spatial details**
- **Thus, can compress more aggressively in U & V than in Y**



Original Y U V

Interchangeability of Color Spaces

- **One can map back and forth between any two 3D color spaces via matrix multiplication (using an appropriate matrix and its inverse)**

- **For example:**

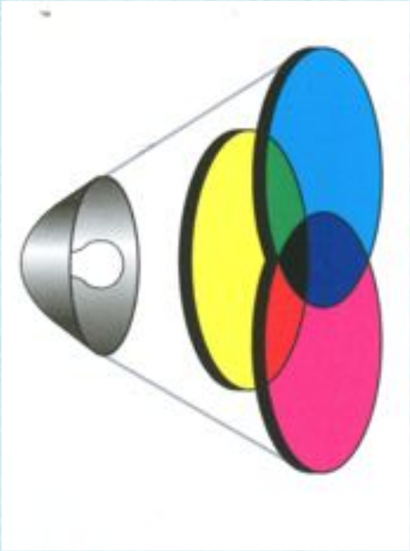
$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ -.14713 & -.28886 & .436 \\ .615 & -.51499 & -.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- **Aside: note how important the Green channel is for the details in Y, as well as how unimportant the Blue channel is (for detail)**

Additive vs. Subtractive Color Spaces

- **Additive Color Space:**
  - **Superimposed colored lights (e.g. computer display)**
  - **Add spectra (wavelength by wavelength)**
  - **R + G + B = white**

- **Subtractive Color Space:**
  - **Sequence of color filters (e.g. ink pigments or paint)**
  - **Multiply by all absorption coefficients (wavelength by wavelength)**
  - **R + G + B = black**



Printers (CMYK)

- **Printers use a subtractive color model**

- **Cyan, Magenta, Yellow (CMY) are the three primary colors of the subtractive color model**
- **These inks partially or entirely mask/filter/absorb colors on a white background, reducing the light that would otherwise be reflected**

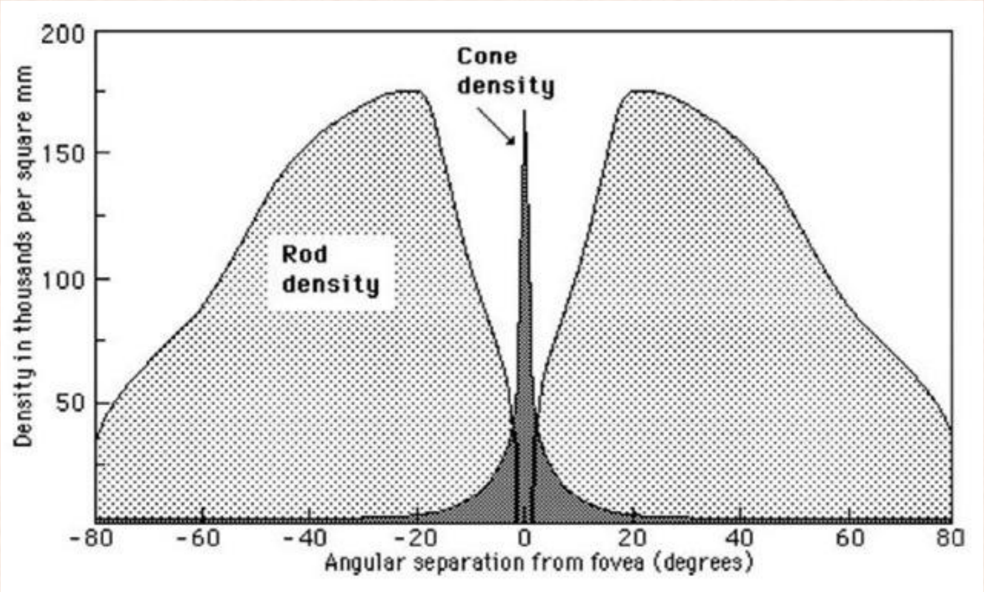
- **Equal mixtures of C, M, Y should (ideally) produce all shades of gray**
- **However, in practice, mixtures of C, M, Y do not give perfect grays**
- **In addition, it’s difficult to get perfect alignment of the 3 inks**

- **Instead, most fine details are printed with the Key color (= K = black)**
- **This also reduces ink bleeding, reduces the time to dry, and saves money on colored ink**

C
M
Y
K

Limited Spatial Resolution

- **Sensors are have a finite size, or area**
- **Thus, there is a limited number of signals per square inch, based on how closely they are packed together**



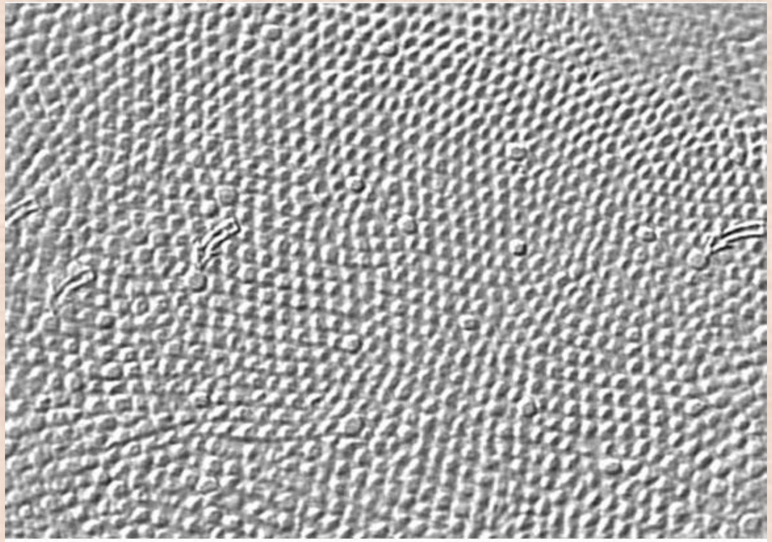


Fig. 13. Tangential section through the human fovea. Larger cones (arrows) are blue cones.

- **The cones are the most densely packed at the center of the retina (the fovea), giving maximum detail for whatever the eye is looking directly at**
- **The rods have nearly a zero density at the fovea, which is why astronomers look out of the "side" of their eye**

Distance Matters

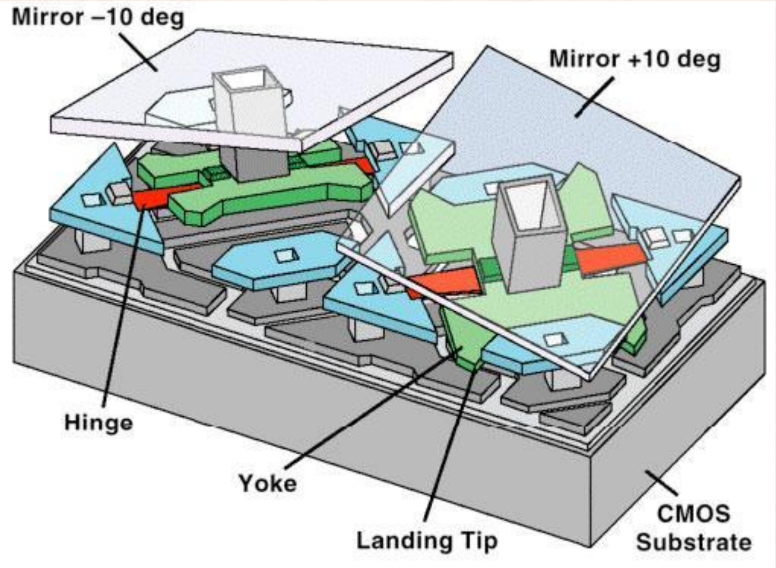
- **Closer/farther away objects project to larger/smaller areas on the cones, meaning more/less cones receive light signals from it**
- **Thus, closer objects can be seen in higher spatial detail than farther away objects**

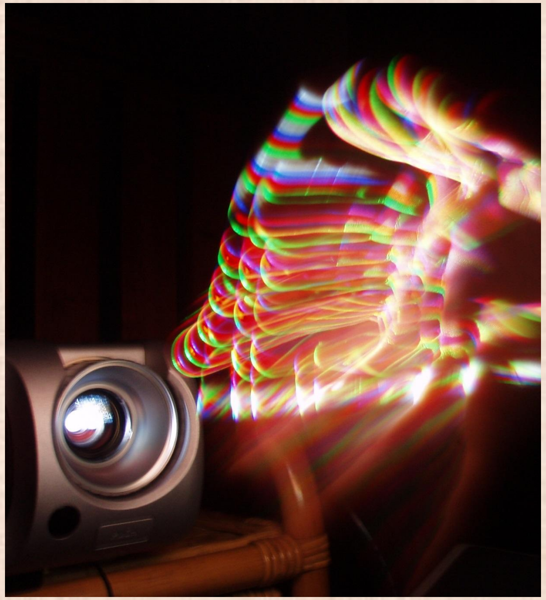
Resolution: 2048x1080
Size: 13.7m diagonal
4.29 dots per inch (dpi)

- **Lower resolution is acceptable for a cinema screen, since viewers sit much farther away from it as compared to a cell phone - with 300+ pixels per inch (ppi)**
- **The number of cones per (image) feature is comparable between cinema screens and cell phones, given the differing distance of the observer**

Projectors

- **Making large displays for far away viewers is difficult; thus, projectors are very important**

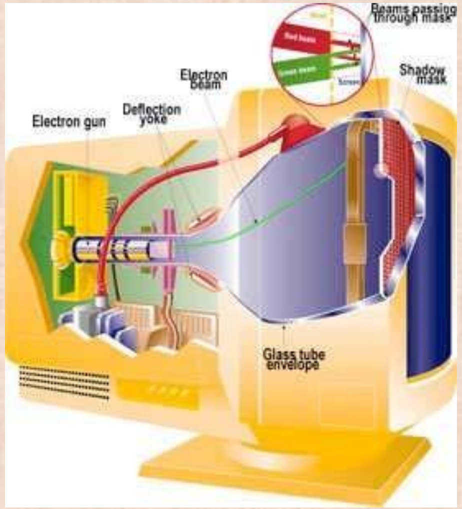




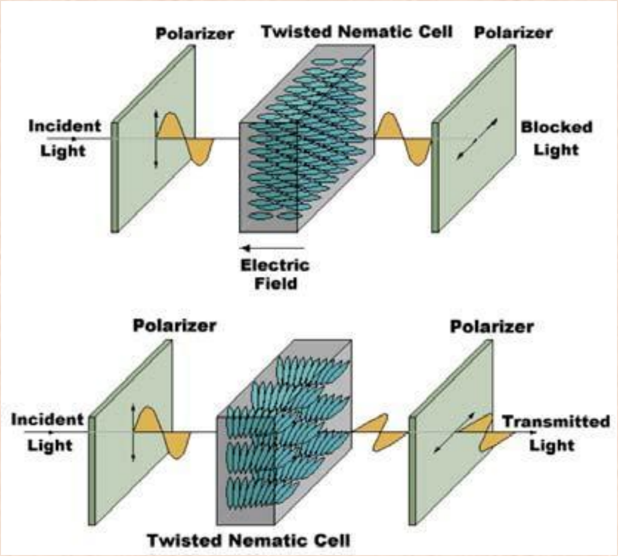
- **A Digital Micro-Mirror Device (DMD) is the core component in Digital Light Processing (DLP) projectors**
- **Each mirror corresponds to one pixel, and has two states; it can either reflect the light into or out of the "pupil" of the projector**
- **Rapidly toggling a mirror between these two states produces brighter/dimmer light, controlled by the ratio of on-time to off-time**

Display Technology

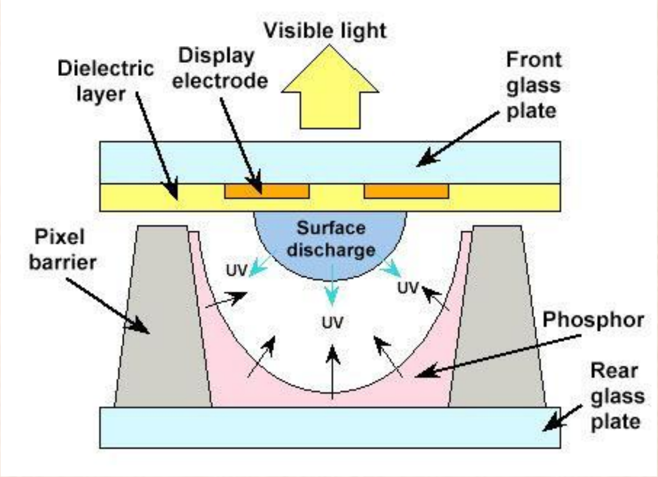
- The closer one sits to a display, the more cones per feature and thus more detail one can see
- Thus, significant efforts have been spent on improving display (spatial) resolution



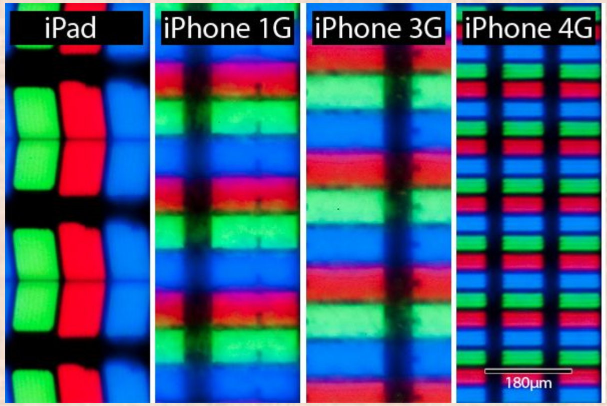
cathode ray tube (CRT)



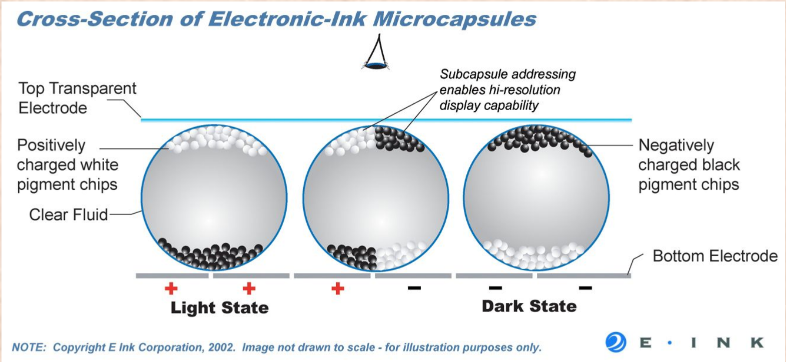
liquid crystal display (LCD)



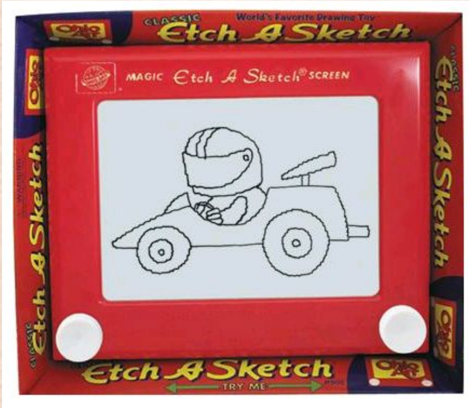
plasma



iPhone/iPad LCD

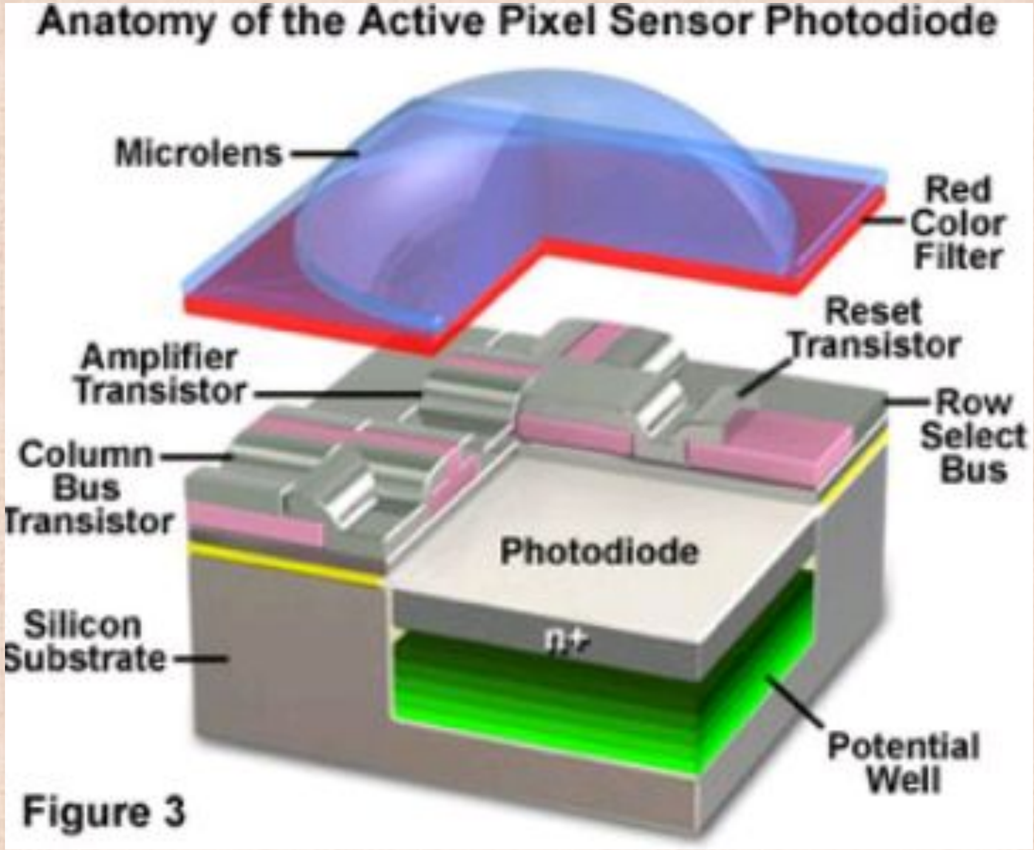


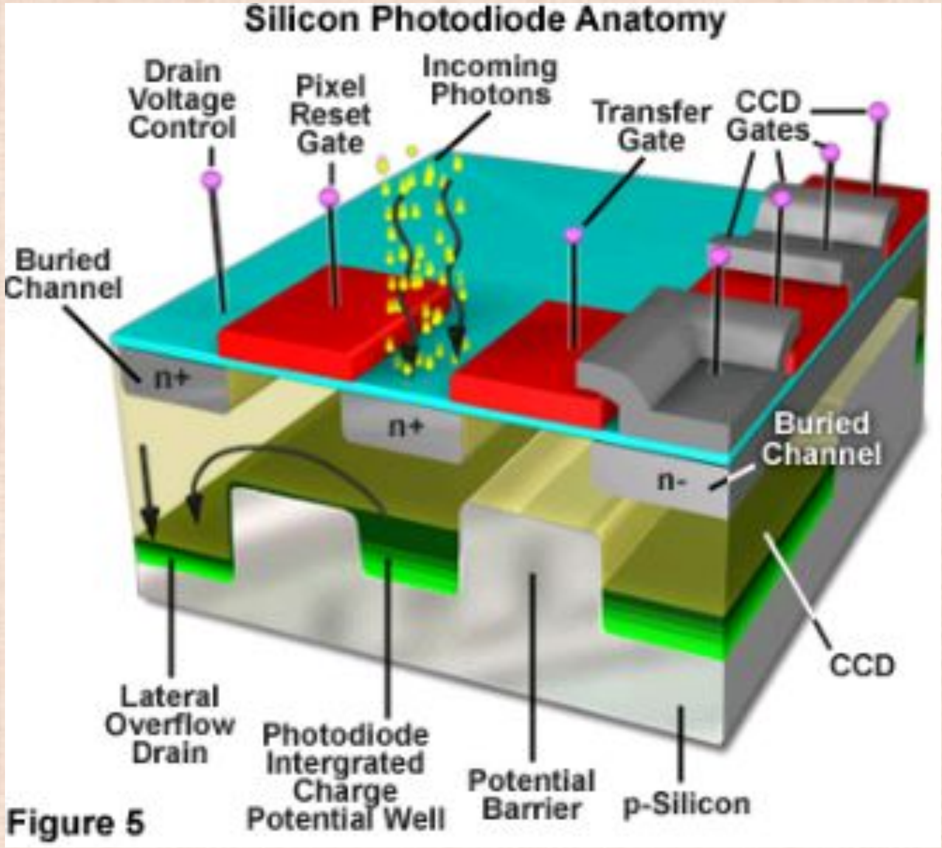
Electronic ink, ebook readers



Camera Resolution

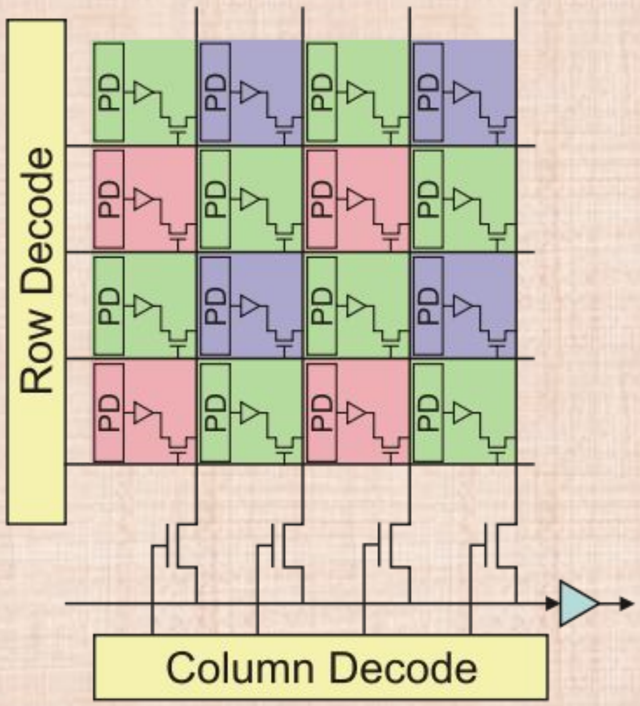
- Camera pixels use the photelectric effect to generate an election when hit by a photon (with some efficiency/probability)
- They are quite complex and, like cones, take up physical space
- This limits the resolution of what they can capture (just like for cones in the eye)



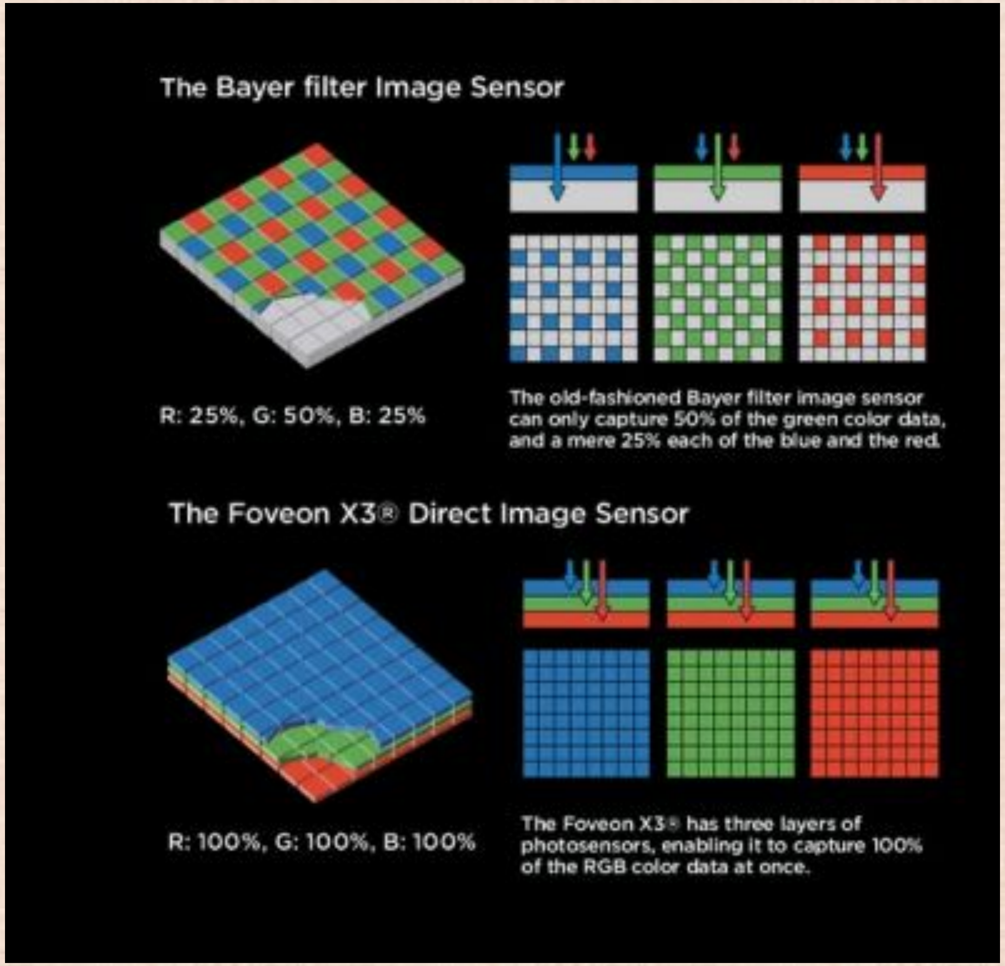


Camera Resolution

- **Each camera sensor records incoming light energy per second (power)**
- **Each captures only one signal (per unit time) for its entire 2D spatial area**
- **Color filters are used to limit incident light to a particular color (so the same sensor can be used for every color)**



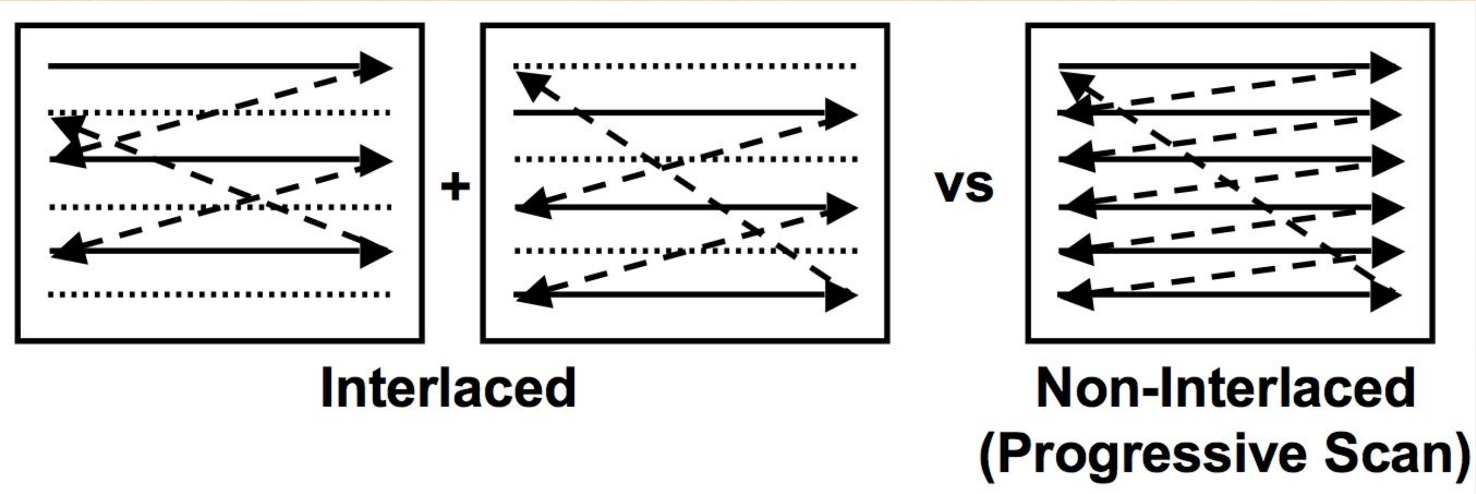
Note the doubled amount of Green sensors, due to that color's importance in capturing spatial details



Temporal Resolution

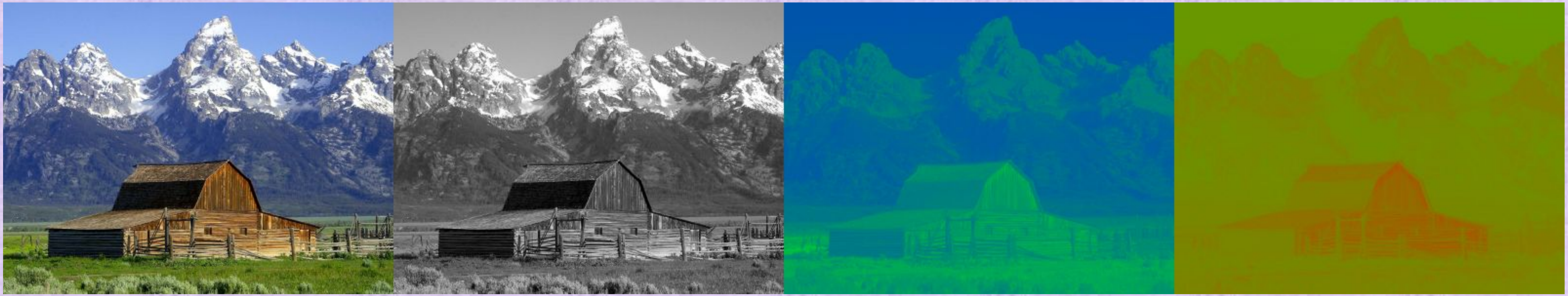
- For <u>moving images</u> (animations), 16 Hz (at a minimum) is needed for humans to *not* interpret them as a series of still pictures
  - Movies are recorded at 24 frames per second
  - TV broadcasts at 30 frames per second

- <u>Flicker fusion threshold:</u> frequency at which an intermittent light stimulus appears to be steady to the observer
- Even though motion may seem to be continuous at 24-30 frames per second, the <u>brightness</u> may still seem to flicker
  - Movies are refreshed at 48 or 72 Hz (with each frame projected 2 or 3 times)
  - Computer monitors refresh at 60-80 Hz (or more) independent of what is being displayed
  - TVs (used to) use interlacing to approximate 60 Hz, showing half of each frame at a time



Brightness (Luminance)

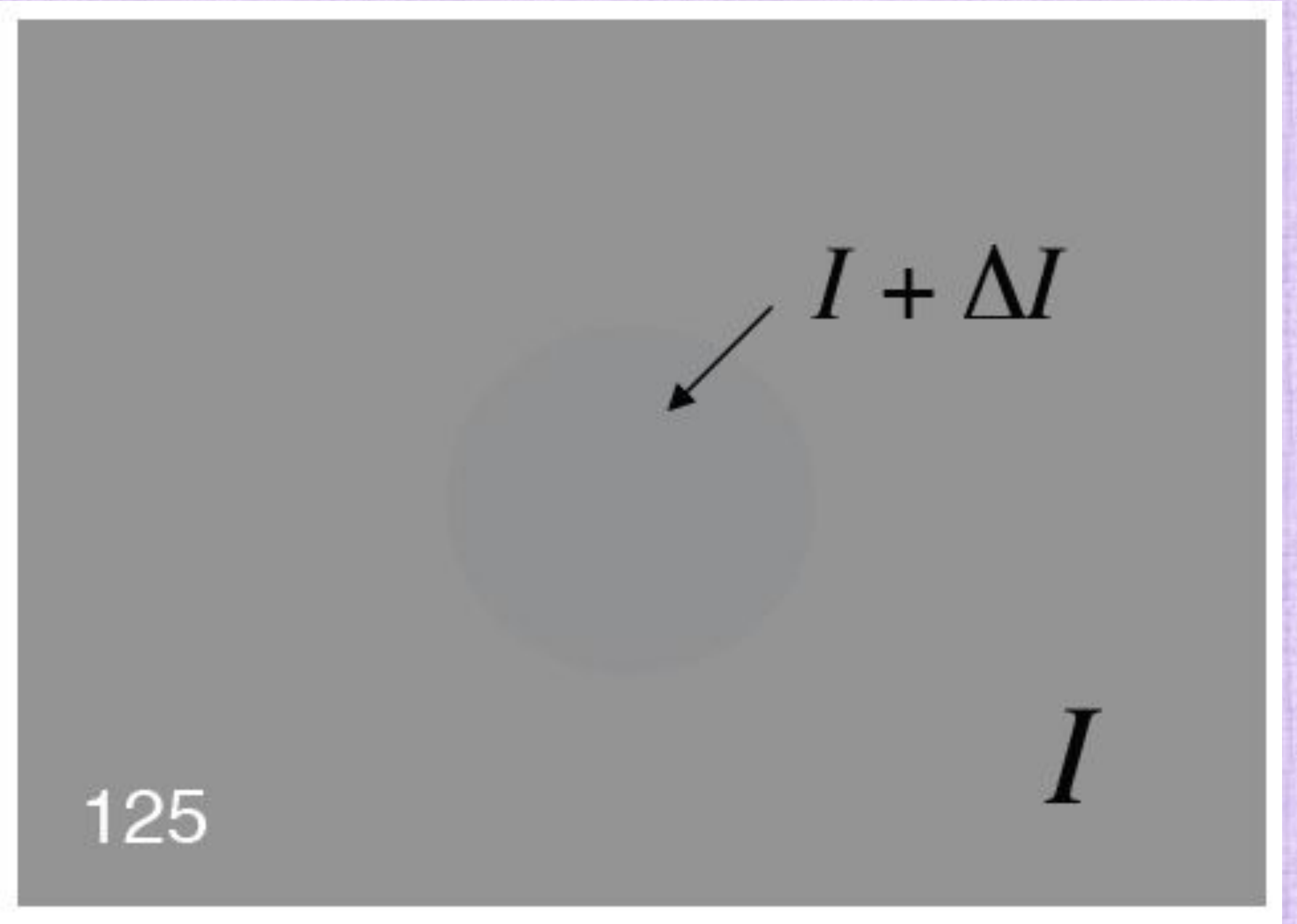
- **The human eye is much more sensitive to spatial variations in brightness (gray scale) than to spatial variations in color**
- **The three images on the right add together to give the image on the left**
- **Notice which of the three images on the right has the most spatial details!**



Original Image = Greyscale + Color 1 + Color 2

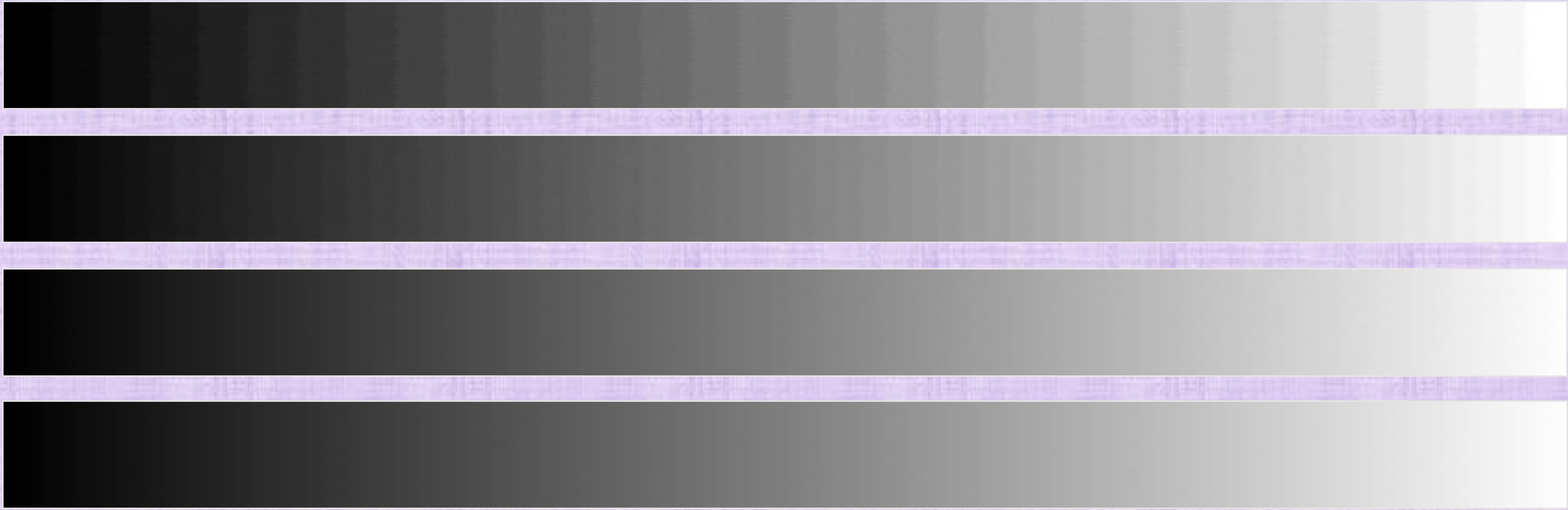
Brightness Discrimination Experiment

- **Changing the brightness (intensity) of the circle by 1 to 2% makes it just noticeable to most people**



Discretizing Brightness

- **Since our eye can see small brightness changes, we need many levels for brightness**
- **Otherwise, changing brightness by the smallest amount looks discontinuous**
- **Thus, we typically use 256 levels for brightness**
- **That is, we store R, G, B each ranging from 0 to 255**
- **High Dynamic Range (HDR) image formats use an even larger range than 0-255**



32 levels

64 levels

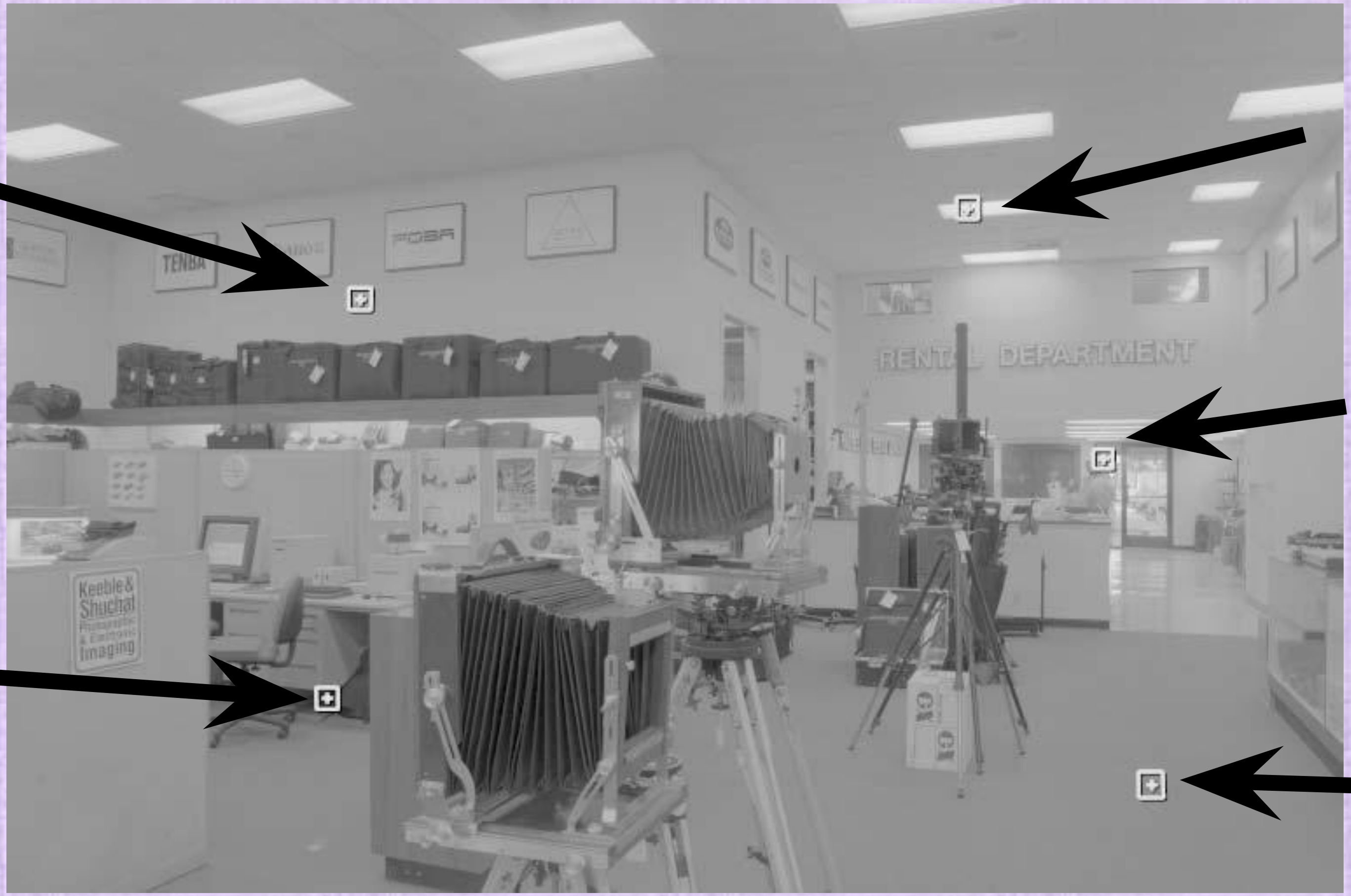
128 levels

256 levels

Dynamic Range

- **World:**
  - **Possible: 100,000,000,000:1 (from the sun to pitch black)**
  - **Typical Real World Scenes: 100,000:1**
- **Human Eye:**
  - **Static: 100:1**
  - **Dynamic: 1,000,000:1 (as the eye moves, it adaptively adjusts exposure by changing the pupil size)**
- **Media:**
  - **Newsprint: 10:1**
  - **Glossy print: 60:1**
  - **Samsung F2370H LCD monitor: static 3,000:1, dynamic 150,000:1**
    - **<u>Static contrast ratio</u> is the luminance ratio between the brightest white and darkest black within a *single* image**
    - **<u>Dynamic contrast ratio</u> is the luminance ratio between brightest white possible (on any image) and the darkest black possible (on any image) on the same device**
  - **The contrast ratio in a TV monitor specification is measured in a dark room. In normal office lighting conditions, the effective contrast ratio drops from 3,000:1 to less than 200:1**

The Real World has High Dynamic Range



The relative irradiance values of the marked pixels

The Real World has High Dynamic Range

- **16 photographs of the Stanford Memorial Church taken at 1-stop increments from 30s to 1/1000s**
- **No single image captures everything desirable in both the darkest and the brightest regions (some pixels are over-saturated and others have no signal at all)**

From Debevec and Malik, High Dynamic Range Photographs

Tone Mapping

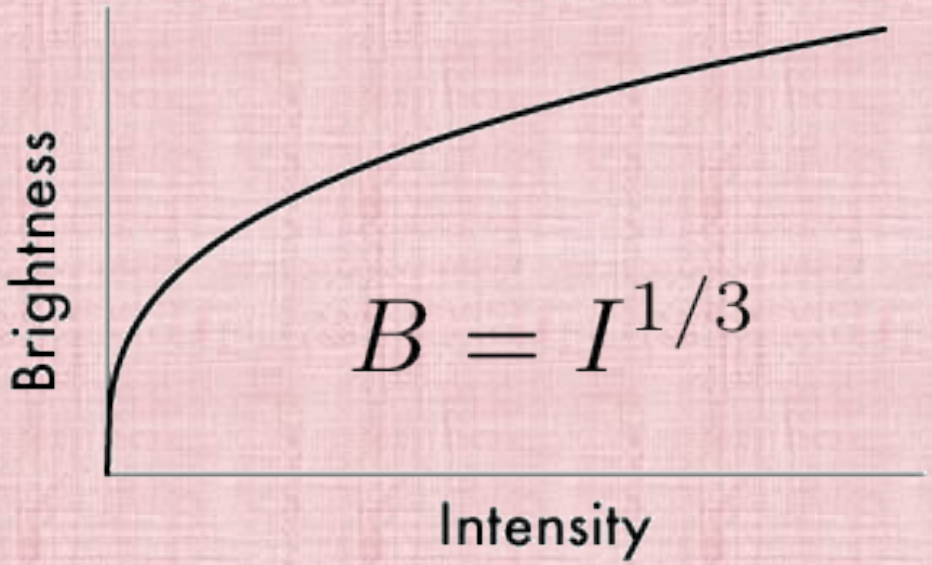
- "Compositing" <u>all</u> the information from <u>all</u> the images gives a result with a <u>High Dynamic Range</u> (i.e., 0-X with X >> 255)
- But that range is too large for the standard image format (i.e., since X > 255)
- Solution #1: Linearly rescale/compress the values so that X=255
  - Small intensity differences are <u>quantized</u> (a range of values map to the same integer), and relative differences (and details) are lost
- Solution #2: Logarithmic map to rescale/compress
  - Information is still quantized, but in a more forgiving way exploiting human's "perceptual space" (once again)
- Solution #3: Other approaches…
  - E.g., Local operators - map each pixel value based on surrounding pixel values (human vision is sensitive to *local* contrast)

Human Perception of Intensities

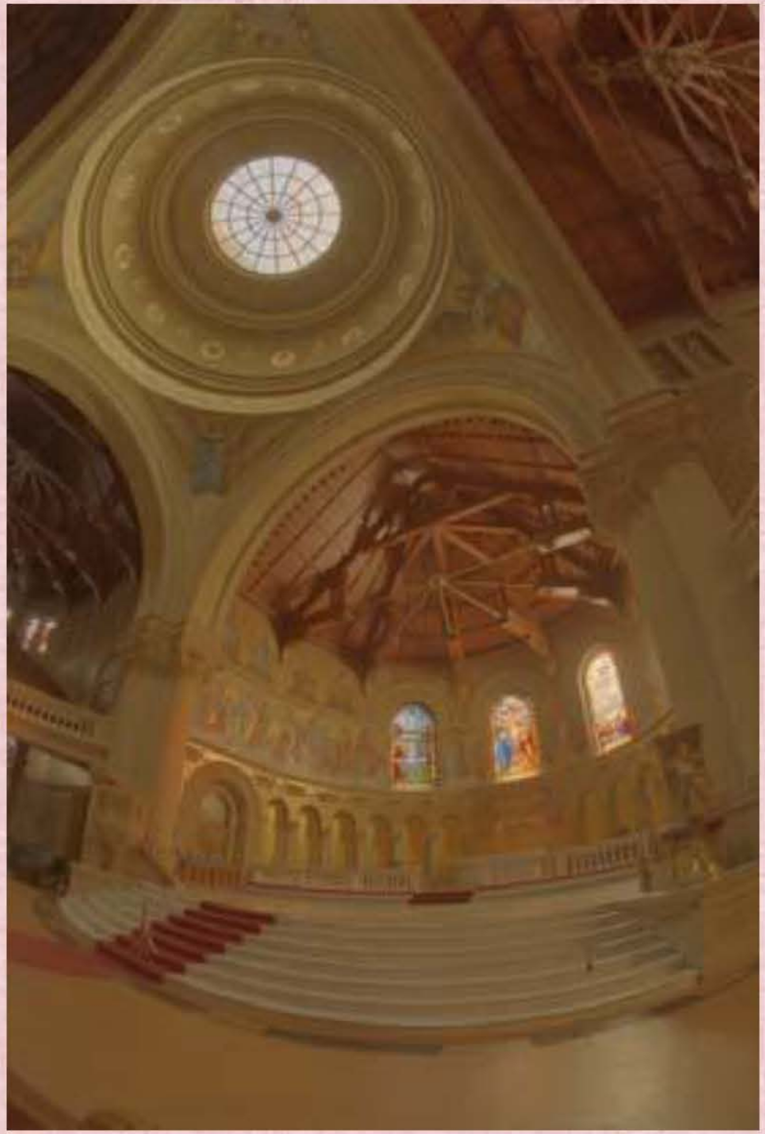
- Brightness intensity differences are better perceived by humans at lower (as opposed to higher) light intensities (a 1-2% difference is smaller at lower intensities)
- Logarithmic compression uses more of the display's brightness resolution for the more important lower intensities in the image (and thus less for the higher intensities)
- This gives less quantization in the lower intensities of the image (than in the higher intensities), and is thus more optimal for human consumption

$$S = I^p$$

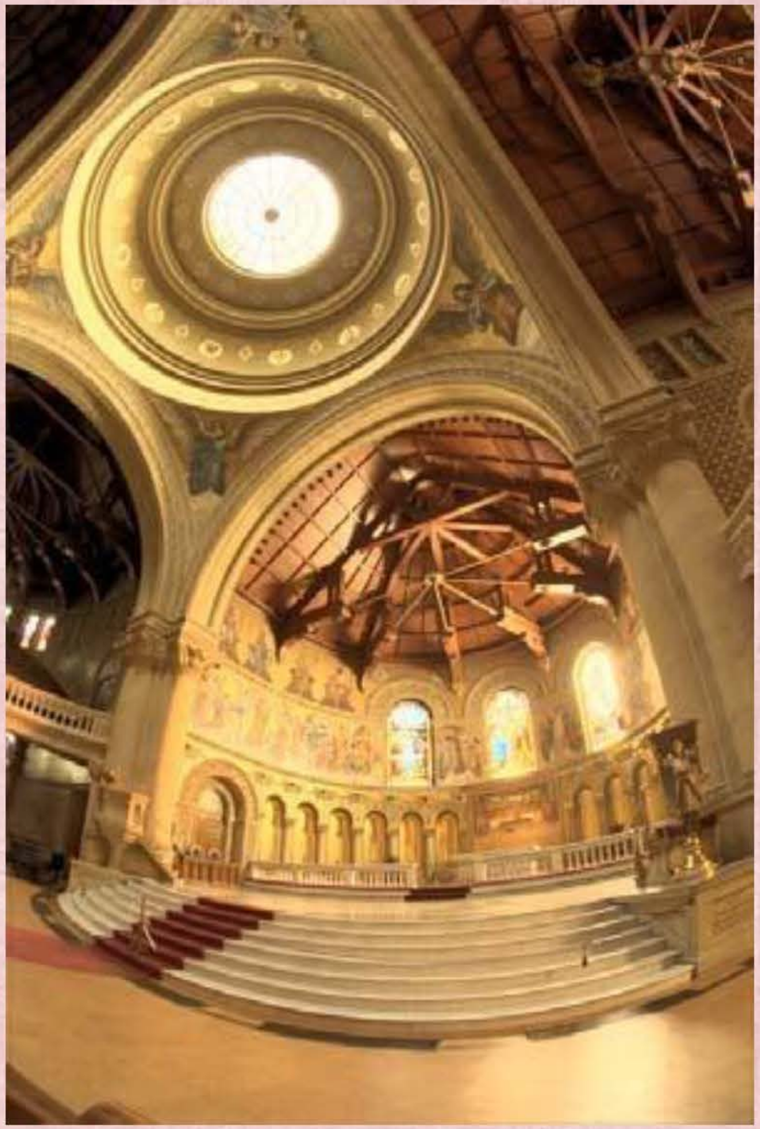
Sense	Exponent
Brightness	0.33
Loudness	0.60
Length	1.00
Heaviness	1.45



Linear vs. Logarithmic Compression



Linear



Logarithmic

Gamma Encoding and Correction

- **Maximize the use of the information relative to human perception**
- **More bits are allocated to the lower intensity (darker) regions of the image than to the higher intensity (lighter) regions**
- **<u>Gamma correction</u> is applied to the <u>gamma encoded</u> images to convert them back to the original scene brightness/luminance**